Contract #FA9550-10-C-0151

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
11-04-2011	Final Report	15 Jun 10 - 14 Mar 11

4. TITLE AND SUBTITLE

Phase I. Final Technical Report and Phase I Final Summary Report
Coupled Cluster Methods for Multi-Reference Applications

5a. CONTRACT NUMBER
FA9550-10-C-0151

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Rodney J. Bartlett, University of Florida

Mark Ponton, ACES QC, LC

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

ACES QC, LC
1421 NW 47th Terrace
Gainesville, Fl 32605-4564

8. PERFORMING ORGANIZATION REPORT NUMBER
AFSTTRI 03

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/PKC
875 North Randolph Street
Arlington, VA 22203-1768

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Report developed under STTR contract for AF09-BT40. In Phase I we have studied several different approaches to the multi-reference coupled-cluster (MR-CC) problem, all the existing ones and introduced three new ones. The latter are the Tailored CC, (TCCSD(T)), the multi-reference double ionized and double electron attached (MR-DI/DA-CC), and a new adapted method,(@CC) which guarantees accurate results for single or multi-reverence problems and for ground or excited states. We believe these offer better alternatives for wide spread application to molecular structure and spectra. All have been compared on examples including the auto-isomerization of cyclobutadiene from its rectangular form through its square transition state, back to its other rectangular form. This example is notable as it has one of the largest numerical effects seen for multi-reference character measured as the difference between the single-reference results given by CCSD(T) and MR-CC. The correct answer for the barrier is expected to be <10 kcal/mol yet CCSD(T) gives 18 kcal/mol. TCCSD(T) gives 7 kcal/mol and MR-DI-CCSD gives 10 kcal/mol. We also made applications of the latter to several different potential energy curves to demonstrate that unlike SR-CC, MR-DI-CCSD provides correct separation all the way to the separated atom limit in a purely symmetry adapted way. This opens a new, attractive route to most MR problems.

15. SUBJECT TERMS
STTR Phase I Final Summary Report

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Rodney J. Bartlett
U	U	U	U		19b. TELEPHONE NUMBER *(include area code)* 352-392-6974

PHASE I FINAL REPORT AND FINAL TECHNICAL REPORT. COUPLED CLUSTER METHODS FOR MULTI-REFERENCE APPLICATIONS

A. Relevance to DoD

Within the DoD the development and application of 'predictive' *ab initio* quantum chemistry has a very high priority. This is because often the best route toward hard-to-get information about atoms, molecules, clusters, and solids is through the direct solution of the Schrödinger equation (or Dirac equation for heavier elements) that describes the electrons in such systems. With sufficient accuracy, such results will offer a prediction in the absence of experiment, will be able to describe molecules in extreme situations not amendable to experiment, or will offer another voice to aid in the interpretation of experiment. This interplay between theory and experiment is on prominent display at the AFOSR reviews and is a powerful testament to this synergism. To mention just a few of the topics addressed are the determination of rate constants for atmospheric species, cross sections for collision induced vibrational excitations for the detection and identification of species from rocket plumes, combustion products, the design of new materials, which involves studies of novel clusters whose structure and spectra requires coordinate theory applications; and the search for new, highly energetic (HEDM) fuels including those based upon ionic liquids.

The ability to obtain the kinds of results required for molecular structure, spectra, energetics, and other properties and to have confidence that they are accurate to within expected error bars, requires two elements: new theory developments in how to better solve the Schrödinger equation and their implementation into software. Also, it is far preferable that the software that makes these highly complex calculations can be readily used by non-experts. In connection with the latter, it is also required that the methods be applicable to much larger systems than in the past, which today demands that they be implemented into massively parallel programs like ACES III that can be run at DoD funded HPCMP's.

B. Current Standard is Single-Reference Coupled-cluster Theory

Today, the reference *ab initio* approach for the largest number of molecular problems is usually single-reference coupled-cluster (CC) theory, built on the ansatz that the correlated wavefunction is $\Psi=\exp(T)\Phi_0$, with Φ_0 a single determinant (independent particle) approximation [1]. This is due to the fact that CC theory normally offers the most rapid convergence to the exact full CI (exact) results for molecules, as shown in Fig. 1. CCSDTQP would means respectively single, double, triple, quadruple, and pentuple excitations, which means $T=T_1+T_2+T_3+T_4+T_5$. CCSD means limit T to only T_1+T_2. CCSD(YT) adds anon-iterative inclusion of triple excitations. (See Fig. 1.)

20110622003

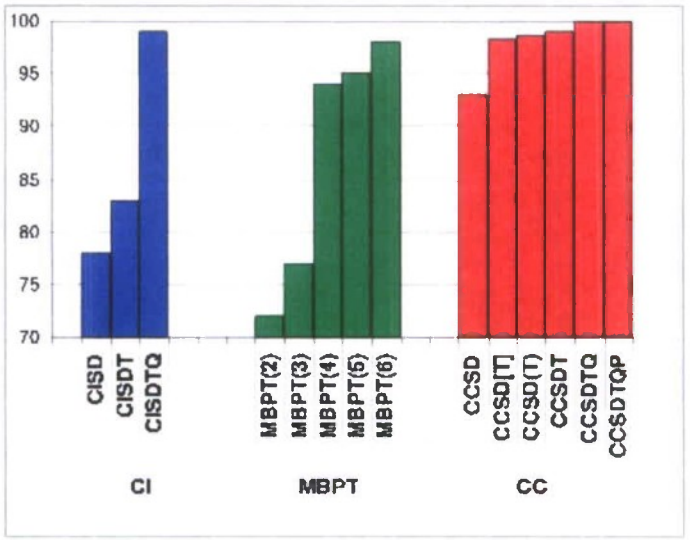


Fig. 1

When SR-CC fails to offer a good approximation it is usually because of the unsuitability of the single determinant reference. That dictates the demand for a multi-reference approach in CC theory *that will be as easy to use as are the current SR-CC methods.* Though several MR-CC methods exist, none offer the ease of application of SR-CC and, consequently, would not be suited to wide-range application to molecules, particularly by non-expert users.

In most respects the single-reference CC method is ideal from both a practical and theoretical viewpoint. For the first, SR-CC methods have the innate simplicity that the only decisions required to use them are the basis set, like choosing cc-pVTZ, and the level of correlation like CCSD(T). As for the second, it rapidly exhausts all the 'dynamic' correlation required to provide predictive results for molecular structure and spectra. Because of the wavefunctions' exponential form, $\Psi=\exp(T)\Phi_0$, for any approximation to T satisfies the fundamental property of size-extensively (linked diagram structure). This also means at the separated limit $\Psi_{AB}=\exp(T_A+T_B)\Phi_A\,\Phi_B=\exp(T_A)\Phi_A\exp(T_B)\Phi_B$. This essential property cannot be achieved without the exponential ansatz of CC theory, which is the primary failing of CI approximations [2,3].

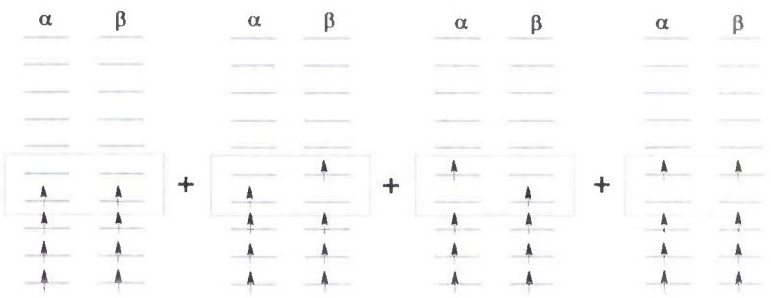
Since SR-CC has to converge to the full CI, they only fail when it is not feasible to describe a problem at a reasonable level like CCSD (o^2v^4), or CCSD(T) (o^3v^4), but might require CCSDTQ (o^4v^6), and higher. (o indicates the number of occupied orbitals and v the number of virtual orbitals.) Each time a new level of excitation is added, the scaling of the calculation in the number of occupied and virtual orbitals rapidly increases, as indicated. There are many ways to reduce the scaling [4], but for the purpose of the present work, we are interested in improving upon the SR-CC by introducing additional effects of MR character.

Furthermore, in many respects the straight-forward SR-CC route adds too many higher excitations into the problem, of which only a subset might be numerically important. To make

the transition to MR-CC we address the issue of an 'interacting' space. An interacting space means we consider the determinants that mix strongly with an initial set of single and double excitations, regardless of their excitation level. This is meant to add the numerically most important determinants to the wavefunction while being less constrained by the theory to add the other less important ones.

C. Nature of Multireference Problem

To give an example of the intent, consider the four determinants:



When the two spatial orbitals in the box, call them I and A, are close in energy or otherwise strongly interacting, then these four determinants composed of spin-orbitals Iα and Iβ and Aα and Aβ are all expected to be important in the wavefunction. That means they should have comparatively large weights in the full CI (exact) solution for the problem. To correctly separate the H_2 single bond, without symmetry breaking as in UHF, these four determinants are required. For this case, I is the σ_g bonding orbital and A would be the σ_u antibonding orbital, which are exactly degenerate at complete separation.

This kind of mixing of four determinants is sometimes called left-right correlation or more commonly, *non-dynamic*, since it is not simply keeping electrons apart, but instead reflects some basic aspects of the electronic structure as this example encounters quasi-degeneracy in bond breaking. Both dynamic and non-dynamic correlation have to be properly described in molecules, and including the latter, gives rise to what are called multi-reference problems.

The generalization to a MR in CI is straight-forward, as it just entails adding typically all the single and double excitations that can be made from these four determinants into the variational CI wavefunction determination: that would be the 'interacting' space. That expands the important determinants in the MR-CI and that helps to obtain better approximations for some few electron problems. But the MR-CI theory remains fatally flawed because it is not size-extensive.

To the contrary, the generalization of SR-CC to the MR is not straight-forward because the exponential ansatz has to be maintained if the MR-CC is to be size-extensive and benefit from all its essential advantages. There are at least a half-dozen serious attempts in this direction that go under names like Fock Space, valence universal (VU) [5,6], Hilbert space, state universal(SU) [7,8], Hilbert space state specific (SS) [9,10], etc. but *all* have formal and computational limitations. They also suffer from miscellaneous other problems such as either not being able to obtain properties like forces, a necessity for useful applications to molecules, or transition moments for excitation energies, or density matrices for ground and excited states, or higher-order properties like NMR and NLO. To the contrary, SR-CC and its EOM-CC generalizations for excited states do all the above [1-3], and do so in a very convenient 'black-box' way that anyone can apply. For the purposes of the present work, we are interested in improving upon the SR-CC by introducing the additional effects of MR character.

A single reference starting point would build upon the first of the four determinants shown, relegating the other three to the complementary, orthogonal Q space, being introduced as single and double excitations in CCSD. That constrains their coefficients to be those obtained from the CCSD wavefunction equations. The exact energy comes from knowing only the single and double excitation coefficients, but those values arise from the full CI, where the triples and higher excitations have contributed to their values. The SR-CC calculation expects that this single determinant will be a reasonable starting point to rapidly converge to the full CI values as other excitations are added into the problem. However, if the weights for these other determinants are not able to grow as large as they need to be with a reasonable level of excitation to reproduce accurate solutions, then there is a residual single reference bias. This is a more general definition of what is meant by a multi-reference CC problem. Since the SR-CC will produce the full CI as a function of excitation level fairly quickly, a method like CCSD(T) or CCSDT will often provide highly accurate results. But if higher excitations like quadruples in CCSDTQ are important to reach the full CI values, then a better route might be to incorporate all four of the determinants shown into a MR-CC wavefunction in a way that their weights are not dictated by the inclusion of higher and higher excitations. Then we would hope to achieve more rapid convergence, better answers, or less expensive calculations than by simply adding higher excitations to the SR-CC.

As more orbitals begin to strongly interact, then instead of a two-orbital, two-electron problem as illustrated here, we would need to have several 'active' orbitals and, consequently, several determinants being highly weighted. MR-CC methods thus become far more difficult and time-consuming when the number of active orbitals increases. Nearly all genuine MR-CC applications to date are only for two-orbital, two-electron problems. Yet even the N_2 molecule requires a minimum of eight orbitals and octuple excitations among them to correctly separate its triple bond. If we consider a problem like Cr_2, which is said to have a hextuple bond, the difficulties are apparent.

D. New Philosophy for Mutlireference Coupled-Cluster Theory

The straight-forward route toward MR-CC is the state-universal Hilbert space method. It uses the generalized exponential form $\Sigma_\mu \exp(T_\mu)\Phi_\mu$ to define the MR-CC wavefunction. There is a separate T_μ for each of the four Φ_μ in our four determinant example. In fact, the four T_μ are redundant, since some of the same excitations can be generated from more than one Φ_μ. When this ansatz is inserted into the Schrödinger equation it leads to a complicated problem that is meant to describe four states simultaneously. For a problem where all four states are quasi-degenerate, this might be the best route, but in practice, some states like the ground state will be separated from others until maybe bond breaking occurs. Consequently, this state-universal (SU-CC) approach is beset with intruder states. That means that at some geometry on a PES another state besides the four of interest, inserts itself into the problem. This will typically cause a singularity in the equations which prohibits any states to be obtained. General reference space choices help and the Li-Paldus C-conditions, but intruders are still a major problem.

Attempts to fix this have led to the State-Specific (SS) MR-CC formulations, the BW [9] and Mk variant [10]. This in simplest terms is a kind of partitioning of the multi-state SU problem into trying to extract only one state at a time. However, this kind of partitioning can be made in several different ways causing residual problems in these formalisms. The BW (Brillouin-Wigner) variant is not size-extensive, and that is the rational for all CC methods [1]. The Mk (Mukherjee) variant is, but it does not satisfy the projected Schrödinger equation so it does not converge to the full CI for a given basis. Also it is not invariant to choices of active orbitals, which means very different answers can be obtained with a different selection. It also has shown substantial convergence issues. The SS method of Hanrath is another option that is not fully size-extensive but fixes some of the other limitations mentioned. It currently does not offer any properties.

Paradoxically, because of how the theory is constructed [2], unlike a MR-CI, no MR-CC method properly reduces to the SR-CC in the absence of degeneracy. This means for a given effort, while the MR-CC will describe the quasi-degenerate aspect very well, it will perform more poorly than SR-CC for the rest of the problem. (Numerical examples were shown in our Phase I proposal.) We consider this to be a major deficiency that needs to be corrected in MR-CC theories. SR-CC should be a special case so that the need for a MR calculation can be assessed beforehand.

In the following, we take a very different route toward the MR-CC problem. First, we recognize that the exponential wavefunction, $\exp(T)\Phi_0$, which is responsible for size-extensivity, is the ideal way to treat the global, *dynamic correlation* in a molecule. However, it is not the best way to treat the *non-dynamic* part which is the cause of MR problems. This part of the problem reflects the related issue of size-intensivity, ie the

energy differences between two size-extensive solutions. It is size-intensivity that is an issue for the energy differences in excited states, eg, The EOM-CC is size-intensive, as the excitation energy for (AB)*→A* + B or A + B* is obtained. When one attemtps to apply an exponential operator to a MR space, as in the SU approach, one is asking a great deal of the operator since it has to account for both dynamic and non-dynamic correlation. Instead, separating the two effects to some degree, allows the best use of the theory for each part, to provide highly attractive alternative MR-CC methods which we think will offer notable improvements in accuracy and/or ease of use.

E. Technical Objectives

The objectives of our work include the following:

(1) Provide general purpose, multi-reference coupled-cluster (CC) analogs to augment the powerful, single reference CC theory that now defines the standard of accuracy in molecular electronic structure theory;

(2) Implement such MR-CC methods into the massively parallel ACES III program system for facile application to problems of interest to DoD and others.

(3) Demonstrate improvement over the MR-CC methods that currently exist in accuracy.

(4) Enable MR-CC to be routinely applied as a potential 'black-box' by non-experts.

(5) Develop methods that provide applications to ground and *excited* states.

(6) Demand efficient property evaluation, including analytical gradients for molecules.

F. Phase I Accomplishments

Several accomplishments were made in Phase I that we will build upon in Phase II. We did the following:

- Addressed the theory that underlies the current variants of MR-CC, assessed their pros and cons, and proposed and tested some new ones that promise an improvement over the existing approaches. The detailed discussion of these rather involved MR-CC theories is being presented in an extended (~100 page) review article being finalized for publication in Chemical Reviews [11] (Draft available).
- Introduced two new MRCC methods into ACES III, the 'tailored' (MR)TCCSD(T) [12] method and a new multi-ionization/multi-attached MR-MI/MA-CC approach [2,13] in the double (DI/DA) approximation that promises MR-CC with the computational ease of SR-CC.
- Exploited reference states with different numbers of electrons to open the door to routine MR-CC applications with MR-CC methods.

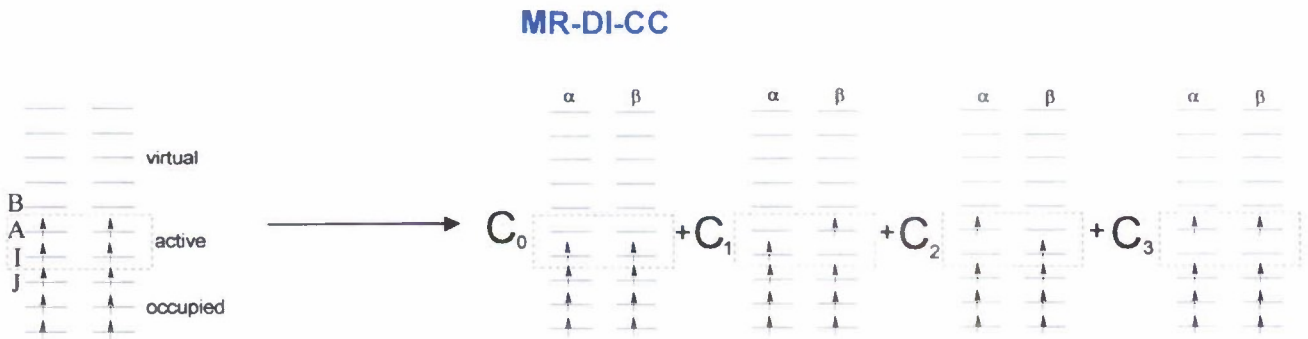
- Made applications to MR problems like PES for twisted ethylene, the isomerization of cyclobutadiene [7,14], the comparative energies .of the mono-cyclic and bicyclic forms of 2,6-pyridynes, and for a variety of bond breaking situations suited to the MR-DI/DA model.
- Invented a new MR 'adaptive' method we term @CC [5]. This method has been assessed using prototype programs we have written and found to be quite powerful. In Phase II we propose to make a general purpose implementation into ACES III. Its objective is to be able to do any state, ground or excited, with an accuracy guaranteed to be below a desired threshold; and it will do this in a way that will be *automatic* for the user.
- Examined the size-extensive EOM-CC to further enable it to fully describe charge-transfer excitations, introducing EOM-CCx as an improvement [15]. The method depends upon the relationship between the MR- valence universal (VU-CC) and EOM-CC, and will have a role in the further development of MR-MI/MA-CC.

Six papers from this STTR have already been published or accepted [2,3,12,13,14,15], with two others [11,16] recently submitted.

- R. J. Bartlett, "The Coupled-cluster Revolution," Mol Phys. **108** (21-23), 2905-2920 (2010).
- R. J. Bartlett, "Coupled-cluster theory and its equation-of-motion extensions," Wiley Inter-Science Reviews, Ed. Walter Thiel, in press.
- D. Lyakh, V. Lotrich, R. J. Bartlett, "The "tailored" CCSD(T) description of the automerization of cyclobutadiene ",Chem. Phys. Letts. **501**, 166-171 (2011).
- M. Musial, A. J. Perera, R. J. Bartlett, "Multi-reference Coupled-cluster Theory: The Easy Way," J. Chem. Phys. 134, 114108/10 (2011).
- D. Lyakh, R. J. Bartlett, "Towards an Adaptive Coupled Cluster Theory," J. Chem. Phys. **133** 244112 (2010).
- M. Musial, R. J. Bartlett, "Charge-Transfer Separability and Size-Extensivity in the Equation-Of-Motion Coupled Cluster Method: EOM-CCx ", J. Chem. Phys., **134**, 034106 (2011).
- D.Lyakh, V. Lotrich, R. J. Bartlett, "Multireference world of chemistry," Chem. Rev., to be published.
- M. Musial and R.J. Bartlett, "Multi-reference Fock space coupled-cluster method in the intermediate Hamiltonian formulation for potential energy surfaces," J. Chem. Phys., to be published.

G. Multi-Reference Double Ionization CC (MR-DI-CC)

We propose a new approach to multi-reference problems that builds upon the use of vacuum states of different numbers of electrons to conveniently and naturally introduce MR character into the wavefunction. We can do this quite generally, with a multi-ionization, multi-attached (MR-MI/MA-CC method. To describe it, consider the double ionization (DI) problem, illustrated below.



$$R^{DI} = \Sigma\ (r_{ij}ji + r_{ijk}{}^{a}a^{\dagger}ijk\ ) = \mathbf{r+rS}$$

$$\mathbf{H^{+2}R=R\omega,\ R=[C,S]}$$

We start from an N+2 electron vacuum as indicated, where H^{+2} means exp(-T^{+2})Hexp(T^{+2}). The T^{+2} is obtained from SR-CC for the N+2 problem. Then we use the $R=\Sigma\ (r_{ij}ji + r_{ijk}{}^{a}a^{\dagger}ijk\ ) = \mathbf{r+rS}$ second-quantized operator to kick out two electrons to return to the N particle problem, our objective. The **S** corresponds to single excitations among the determinants in the MR space. By virtue of doing this, the drawing demonstrates that the four determinants that we expect to be highly weighted in the final solution are naturally introduced into the eigenvalue equation shown above. This equation then assumes the matrix form, $\mathbf{H^{+2}R=R\omega}$, which means that the coefficients in **C** are obtained without any single reference bias, as matrix diagonalization will provide the coefficients in the figure, The ω_k are the eigenvalues for the N particle, quasi-degenerate problem. The scaling after CCSD is only o^4v^2, a fraction o^2/v^2 of the ground state computational effort. This approach conveniently fulfills our objective of determining the important coefficients *without single reference bias*, the essence of MR theory.

Furthermore, this calculation is operationally as easy to do as any SR-CC. In terms of active orbitals, any of the spin-orbitals that are among the new N+2 occupied set can manifest multi-reference or quasi-degenerate effects. Since the *i* and *j* indices cover all such spin orbitals and because results are obtained for any choice of *i,j* from the

solution of the eigenvalue equation, the results apply to the other choices of orbitals occupied in the N+2 elecron vacuum. So in addition the J and A, or I and A, labeled relative to the N particle vacuum, are now included among the new occupied spin-orbitals (i,j,k,...), Other combinations will simply appear as other eigensolutions to the EOM problem. There is no need for a conventional selection of 'active' orbitals among the occupied levels. The one new choice is how many electrons to ionize. If we want to have three of the spatial virtual and occupied orbitals from the original N particle vacuum involved in a MR description, then we set our spin-orbital vacuum at the N+3 electron level to start the process. And these calculations are about as fast as those for the N+2 electron vacuum, because the computational scaling goes down with the number of orbitals being ionized. The generalization to N+4 and N+6 vacuums in ACES III is a major target for Phase II.

It should be understood that the choice of vacuum, which is partly a formal device, does not require using orbitals for the N+2 electron problem, thought that is one choice. Instead we prefer to use the orbitals from the N-particle problem, like the HF solution for the molecule of interest. We simply incorporate the previously unoccupied orbital, A(both spins), in the new N+2 electron vacuum where it is formally occupied to start the procedure. When we go to an N+4 vacuum, we would have both A and B among our new occupied orbitals, allowing them to manifest quasi-degeneracy with any of the prior I, J, K... occupied orbitals.

To demonstrate how well this MR-DI-CC method works, we show results for the highly MR autoisomerization of cyclobutadiene, where we have comparisons to all the proposed MR-CC methods. The results in red are from methods we have introduced.

Table 1. Automerization barrier heights in kcal/mol calculated with different methods in the spherical cc-pVDZ and cc-VTZ basis sets (geometry optimization does not necessarily correspond to the final *ab initio* level presented).

Method	cc-pVDZ, kcal/mol	cc-pVTZ, kcal/mol
CCSD	21.0	23.2
CCSD(T)	15.8	18.3
CR-CCSD(T)	18.3	21.5
ΛCCSD(T)	16.8	19.2
TCCSD	9.4	12.9
TCCSD(T)	4.6	7.0

2D-MRCCSD(T)	6.6	-
SUCCSD(T)	4.8	5.9
BWCCSD(T)(a.c.)	6.1	7.0
BWCCSD(T)(i.c.)	5.7	6.8
MkCCSD(T)	7.8	8.9
RMRCCSD(T)	7.2	9.5
SUCCSD	7.0	8.7
BWCCSD(a.c.)	6.5	7.6
BWCCSD(i.c.)	6.2	7.4
MkCCSD	7.8	9.1
RMRCCSD	10.4	13
MRCISD	7.3	8.4
MRCISD+Q	7.6	8.8
MRAQCC	7.7	8.9
MR-DI-CCSD	8.3	10.7
MR-DI-CCSDT-3	9.00	10.1
MR-DI-CCSDT	8.86	-
SS-EOM-CCSD[+2]	8.3	9.5
Experimental range	1.6–10	

The first four single reference calculations are off by more than a factor of 2. The generalized single reference approaches like our ΛCCSD(T) and the completely renormalized (CR-CCSD(T)) approach of Piecuch do not work. The first MR-CC result is 2D-MRCCSD(T) [7]. The other MR-CC methods that have been applied to this problem, the state-universal, SU-CC [Li, Paldus], the state-specific Brillouin-Wigner (BW) methods [Pittner], and the state-specific MkCCSD method [Evangelista], the reduced multi-reference (RMR) [Li, Paldus], and our MR-AQCC [Szalay, Bartlett] method used by Lischka are also shown, as is our 'tailored'TCCSD [Hino, Bartlett] and TCCSD(T) [12]. Also shown is a SS-EOM-CC [Nooijen]. The answer is not known accurately experimentally.

The MR-DI-CCSD method and its doubly attached variant (MR-DA-CCSD) (the latter uses an N-2 vacuum) also provide very attractive PES for many problems. See the following figures 2-5,

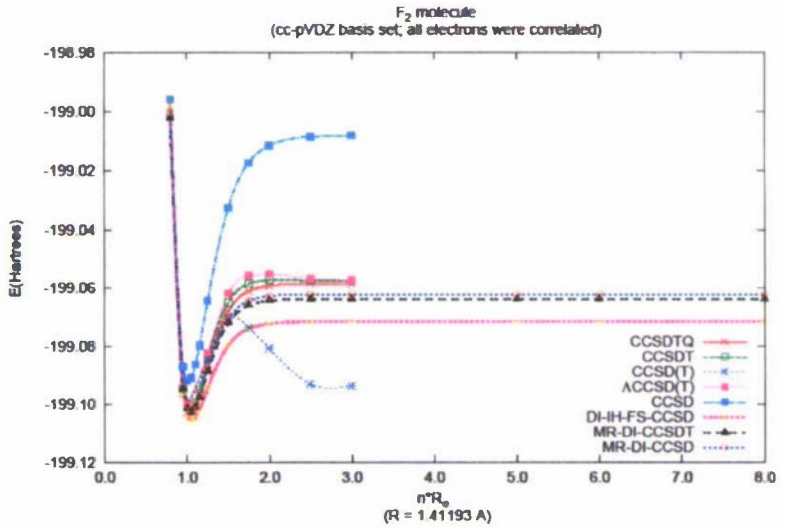


Fig. 2

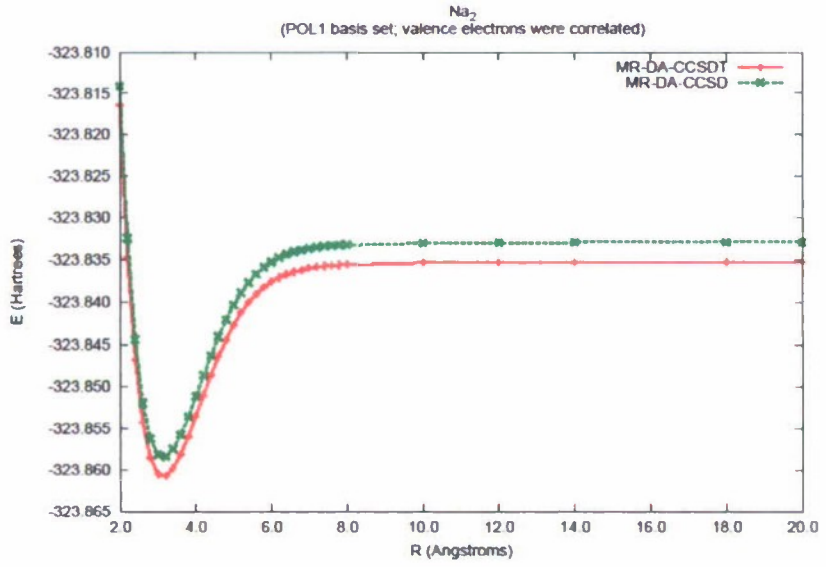


Fig. 3

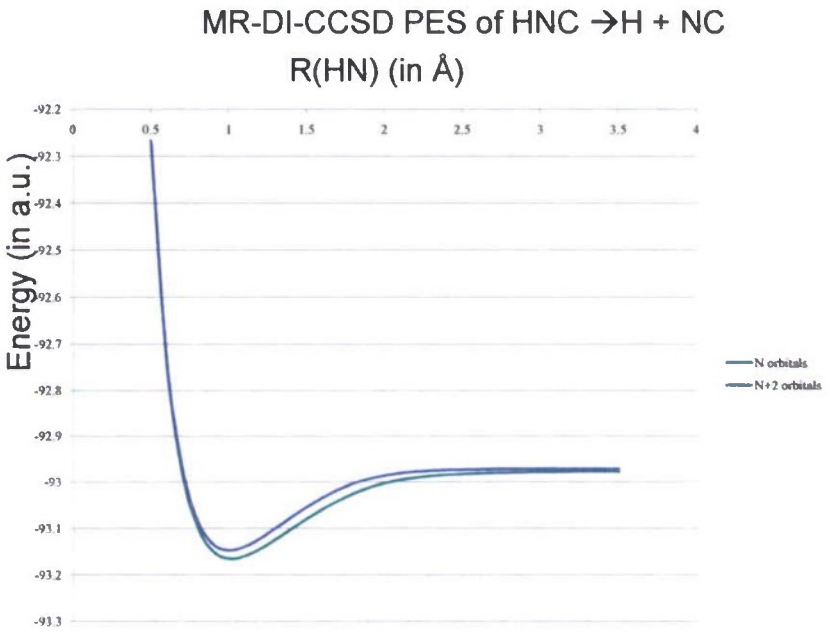

MR-DI-CCSD PES of HNC →H + NC
R(HN) (in Å)
Energy (in a.u.)
-92.2
-92.3
-92.4
-92.5
-92.6
-92.7
-92.8
-92.9
-93
-93.1
-93.2
-93.3
0
0.5
1
1.5
2
2.5
3
3.5
4
N orbitals
N+2 orbitals

Fig. 4

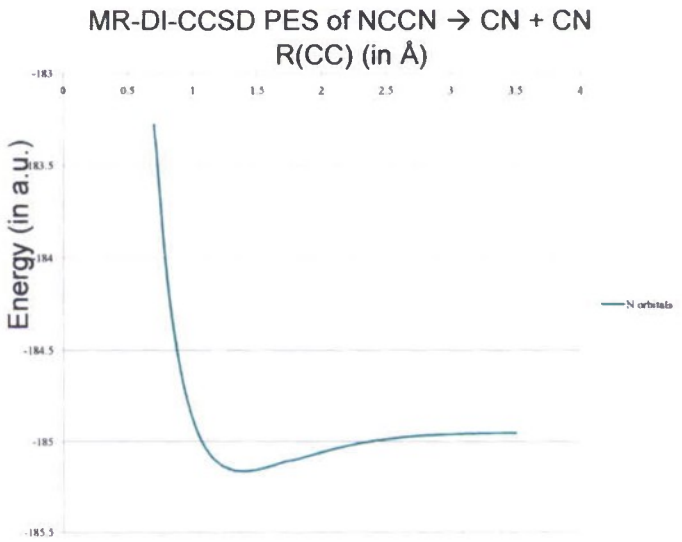

MR-DI-CCSD PES of NCCN → CN + CN
R(CC) (in Å)
Energy (in a.u.)
-183
-183.5
-184
-184.5
-185
-185.5
0
0.5
1
1.5
2
2.5
3
3.5
4
N orbitals

Fig. 5

These curves exploit the fact that either the *formal* double ionization product is two closed-shell anions or two closed cations, but the N particle problem provides the correct description of the radical products.

Bond breaking is always of concern. Here it is clear that all such curves go all the way to the separated limit without any spin-recoupling as occurs in UHF based calculations. Figure 2 shows how well this method works compared to full CCSDT and ΛCCSD(T). The 'gold-standard' CCSD(T) will inevitably turn over as the bond is broken subject to an incorrectly separating RHF reference. ΛCCSD(T is our improvement over CCSD(T). It nicely avoids this, but that method and CCSDT are not easily converged beyond 3x equilibrium, as the figure shows. To the contrary, MR-DI-CCSD has no problems. Figure 6 illustrates the MR-DA-CCSD variant using the N-2 (double cation) vacuum. When there are two curves in the Figure, one uses the N+2 electron orbitals and the other the N electron ones. Being able to use a wide variety of orbitals offers an important degree of flexibility in the method. When there is one curve as for NCCN, it is for the N-particle orbitals meaning the corresponding result for the N+2 did not easily converge. This is why we prefer N-particle orbitals, but they can be HF, KS, GVB, MCSCF, Brueckner, Natural, and others. In Phase II we will explore these choices to find the best among them.

We also show the intermediate Hamiltonian VU-MR-CC result (DI-IH-FS-CCSD) for F_2. Though underestimating the binding energy, this *very* inexpensive calculation ($\sim o^4$) already has the qualitatively correct behavior. VU-MR-CC) is another, component of our effort to obtain the simplest, most effective MR-CC methods possible, and we are uniquely exploring it for PES, while also developing it for excited states and the transition metal multiplet problem.

H. Plans for Future Work

MR-MI/MA-CC

- Generalize the MR-MI/MA-CC to allow for 3-8 orbitals and electrons. Then, just as we showed we can obtain very accurate PES for F_2, Na_2, NCCN, as above, and in the accompanying PacificChem presentation(talk available), We would anticipate similar accuracy for six electron bond breaking as in N_2 and CO, or double bonds as in O_2, C_2, and C_2H_4. At that point, essentially any reaction path will become accessible within a fully spin-adapted method (no symmetry breaking as in UHF based methods.)
- We already have analytical gradients for the standard IP/EA-EOM-CCSD and for excited states, EE-EOM-CCSD in the serial ACES II. We will implement the theory for the new MR-DI/DA-CCSD in Phase II into ACES III, and then its generalizations to the MR-MI/MA variants discussed above. This will enable

automatic identification of critical points on PES, equilibrium geometries, and those for transition states at the MR level.

- At that point we will have as a byproduct all the relevant density and transition matrices and associated properties for these methods, enabling these methods to be used for all properties of molecules.
- Also, as is already done in ACES III for SR-CC at the CCSD and CCSD(T) levels, we will have the tools to obtain the second-derivatives (Hessians) as finite-differences of analytically computed gradients to provide vibrational spectra and to document the nature of the critical point.
- The above will become the first MR-CC method that can be used as routinely as SR-CC is today for all properties of molecules, and it will be available in the massively parallel ACES III program system.
- Just as CC theory generalizes CI by introducing repeated clusters, MR-MI/MA-CC generalizes SR-CC to 2, 3, 4,... electron *quasi-degenerate* variants, a major advance.

In Section B, we showed how MR theory arises in quantum chemistry. SU attempts to create MR-CC methods by starting with these four determinants, and then applying a generalized exponential waveoperator of the form, $\exp(T_\mu)\Phi_\mu$ where there is a different T_μ for each of the four determinants, Φ_μ. In our opinion, this is not the best way to proceed. The *dynamic* correlation is the critical part that depends upon the size of the system, requiring a size-extensive (exponential operator) treatment. The *non-dynamical* quasi-degeneracy correlation is fundamentally size-intensive. This is why the ansatz in this MR theory is uniquely

$\Psi_k = \Sigma(c_\mu \Phi_\mu + \Sigma s_i^a c_\mu \Phi_\mu)\exp(T^{+2})\Phi^{+2}$. The bulk of the wavefunction is the size-extensive $\exp(T^{+2})$ part, while the rest is the intensive part, $\Sigma c_\mu \Phi_\mu$ augmented by the weighted singles. Furthermore, the set of Ψ_k obtained and their ω_k covers all possible two particle quasi-degeneracies, not just those for two particular 'active' orbitals. This is why we see this method as a quasi-degenerate generalization of SR-CC. Generalizations to 3, 4, 5... particles makes the intensive part of the ansatz that much richer, with 3, 4,...electron quasi-degeneracies added. Yet all generalizations can be incorporated in very similar coding. This will offer a powerful, new MR-CC method whose ease of application is exceptional. Further developments in Phase II will address the optimum orbital choices to make these approximations as good as they can be. For example, rather than the previous ones we have used, perhaps GVB or MCSCF orbitals chosen for the intensive part of the wavefunction would be a logical choice.

The other extension we will make to this theory is to allow the MR part of the wavefunction for Ψ_k to define a new cluster operator via the k^{th} set of CI coefficients used in a cluster decomposition to define a new $\exp(X_k)$ operator. This operator will then define a new effective Hamiltonian to use in a second $\exp(T^{+2})\Phi_0^{+2}$ to update T^{+2} . In this

way we allow MR effects to also adjust the amplitudes in T^{+2} to build a state-specific (SS) MR result which can then be repeated until convergence. This procedure should key in to a particular quasi-degenerate k state. The other eigenfunctions then will be byproducts.

Tailored CCSD(T) Method

- The basic idea is to separate the CC amplitude into two parts, a non-dynamic correlation 'valence' part and a dynamic correlation 'core' part, $\exp(T)\Phi_0=\exp(T_{val})\exp(T_{core})\ \Phi_0$. The 'valence' part is dictated by a small number of *valence active orbitals*, indicated by lower case letters, eg CCsdtq, which makes its calculation very fast. CCsdtq scales in terms of active occupied and unoccupied orbitals as $n_{act}^4N_{act}^6$, where $(n_{act}+N_{act})$ will be <10% of the full basis size, saving ~ $(0.1)^{10}$ compared to the CCSDTQ.
- Besides doing a single reference, active orbital CCsdtq... we can equally well do a very fast, all valence active orbital full CI to fix the internal, quasi-degenerate amplitudes from the CI coefficients, $T_{1valt}=C_1$, and $T_{2val}=C_2-C_1^2/2$. Then, the core ones are obtained from standard CCSD, via $\bar{H}=exp(-T_{val})Hexp(T_{val})$, and $\exp(-T_{core})\bar{H}exp(T_{core})\ \Phi_0=E$. With the core$_t$ limited to CCSD, the whole calculation's time is about that for a CCSD calculation.
- Informative results have been reported for the N_2 potential curve, Fig. 6

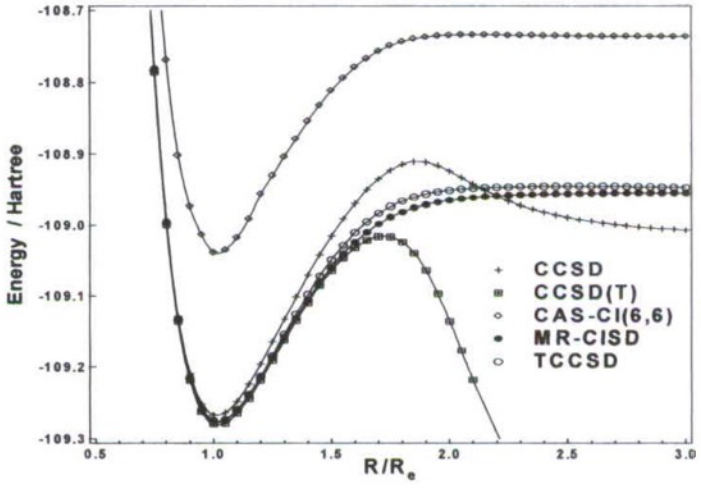


Fig. 6

Where the *first* iteration of TCCSD already gives the exceptional curve shown, correctly breaking the triple bond unlike any of the other current MR-CC efforts.

We also have studied the notorious ozone vibrational frequencies, where only TCCSD gets them right to 0,85%, for the isomerization of cyclobutadiene shown in the Table above, and the relative energies of the two isomers of 2,6 pyridynes where TCCSD(T) is accurate to 0.2 kcal/mol.

- Higher iterations can be accomplished by allowing either the $\hat{H}$ defined by T_{val} in the calculation of T_{core} or build an $\hat{H}$ from T_{core} to modify the T_{val}. The general implementation of CCsdtq requires that certain triple and quadruple CC equations be added to ACES III as a function of the active orbitals. That is another goal for Phase II.
- As a generalized active orbital MR-CC approach, all other properties like analytical gradients will follow just as they do for SR-CC, once the extra terms are coded into ACES III.

We envision the TCCSD and its TCCSD(T) extension to provide a rapid assessment of MR effects in PES with the same computational effort and time as SR-CCSD and CCSD(T).

Adaptive CC (@CC)

- The adaptive CC, is a new idea updating an old one in the field. Namely using determinant selection but in a very different, CC way. The idea is to control the error in any calculation to be below a threshold for any state, ground or excited, and regardless of whether it is single or multi-reference in character [13].
- The main difference between a SRCC which is dictated by systematically adding higher categories for excitations to converge toward the full CI and @CC is that the concept of *an interacting space* is superimposed on the @CC theory making it MR. The interacting space means that for an initial set of single and double excitations, all the important higher excitation determinants with which they interact are allowed to contribute toward the wavefunction without bias toward excitation level.
- To accomplish this, an *internal, dynamically allocated* measure of importance is assigned to every determinant that can mix with an initial set of single and double excitations.
- The @CC calculation is then made subject to including those interacting determinants in the CC calculation. When the important ones change as a function of the state or the geometry, then the dynamic indexing gives a new reference set to continue the process until the errors are reduced below a previously set threshold.
- An important addition needs to be made to ACES III to provide generalized indexing to enable it to use higher categories of excitations routinely. This addition will be made in Phase II, and will also serve methods like CCsdtqph.

The @CC method is based on the coupled cluster ansatz with adjustable amount of variables in it. The internal logic of the method allows a systematic extension of the set of variables until the desired accuracy is achieved in the energy. Thus, this is one of the rare (but highly valuable) approaches with a controlled error. The @CC method is equally suitable for both single- and multireference problems. The only difference is that the number of variables required to converge the energy is higher in the latter case. The @CC approach provides a seamless connection between the single-reference and multireference regions of the PES that will not be mathematically continuous, but will be effectively so because all energies will be correct to within the chosen threshold. The approach can be equally applied to excited electronic states in a state-specific manner (ie, one state at a time).

The @CC method is relatively computationally expensive as will be most methods capable of treating the most general multireference problems. Hence, it is the method of choice when one needs to nail down preliminary or qualitatively correct results. The possibility of a controlled-error description makes the @CC approach a highly-reliable electronic-structure method suitable for almost any problem. The theory of analytical gradients is similar to that used in general-order coupled cluster methods, but will allow for discontinuities below the threshold. Because of its computational requirements the @CC scheme generally requires massively parallel execution, being ideally suited to the ACES III computer package.

At the end of Phase II we will have available three new variants of MR-CC methods, the MR-MI/MA-CC, the TCCSD(T), and the @CC [14] for routine application. All these new methods will be compared to the best currently available reference MR-AQCC results. This method, which we [Szalay, Bartlett] introduced some years ago, is now conveniently available in the COLUMBUS and MOLPRO packages. MR-AQCC results have served as the reference for most SU-CC and SS-MR-CC applications.

We have also identified a number of small test examples of different kinds of MR character, ranging from open-shell transition metal atom multiplets, to the correct description of equivalent valence bond structures , to reaction paths like the Cope rearrangement and Diels-Alder reactions, to the $CH+N_2$ PES studied [17], the $(CuO)_2$ molecule, and the notorious dimer, $(NO)_2$ that has never been adequately described. We will study single, double, triple bonds for bond breaking and even Cr_2. Particular emphasis will be placed on metal containing species and charge-transfer excitations, some of importance to ionic liquids. Of course, ACES III will permit much larger MR-CC applications to be made, and some 'newsworthy' ones will also be studied in Phase II.

It is highly attractive to introduce a diagnostic, which has been considered [54], to indicate when a description of MR character is mandatory to achieve an accurate

result. In an intermediately normalized SR-CC wavefunction, $\Psi=\Phi_0+\chi$. is the correlated part of the wavefunction. When some determinants in χ have large coefficients, then we have some MR indication. ACES writes out the largest values. But this measure can be made more quantitative by studying the norm, $\langle\chi|\chi\rangle$ which should be much less than 1, but is likely to be larger than some threshold when MR effects are not yet included. We propose to develop such a measure from computing this norm for $\exp(T)\Phi_0$ and the other ansätze for the MR examples. We do not consider the T1 diaganostic to offer a meaningful measure because it is orbital dependent, being zero for Brueckner orbitals. However, our T2 form is meaningful. In any event, we plan to investigate the different parts of the norm to define more accurate indicators of MR character.

Excited states built upon MR-CC reference states

Our objectives also include extensive studies of excited states. The basic approach is to follow our EOM-CC method to initiate a new EOM-MRCC method. In a complete description of excited states, obtaining the correct asymptotes for molecular separation is as important as it is for the ground state. In fact, in any kind of response theory like EOM-CC, excited states derive from the ground state description. If the ground state potential energy surface separates correctly, then the EOM-CC description of excited states will have to do likewise as in the Figure. 7

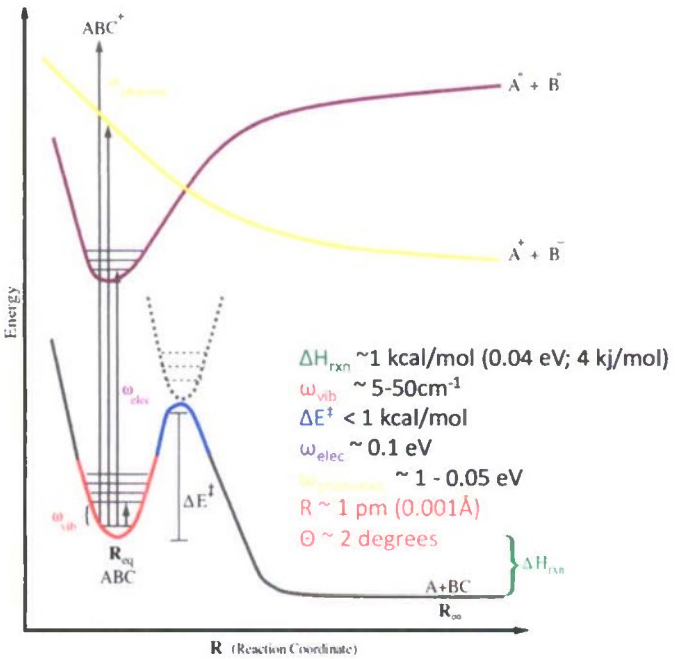


Fig. 7

To make this point, the ground state Schrödinger equation is

$$H\Psi_g=E_g\Psi_g$$

The excited state's equation is

$$H\Psi_X=E_X\Psi_X\ .$$

Once we write the excited state as $\Psi_X=R\Psi_g$, a little manipulation gives the EOM,

$$[H,R]\Psi_g=(E_X-E_g)\Psi_g=\omega_X\ \Psi_g.$$

Once we assume $\Psi_g=\exp(T)\Phi_0$, with [R, T]=0, it follows that

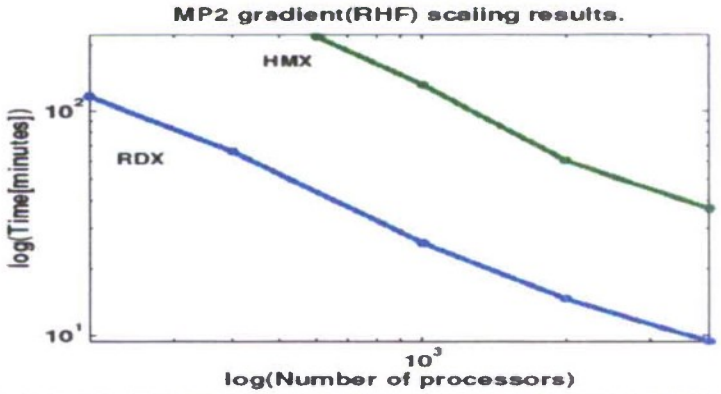
$$\bar{H}\Phi_0=\omega_x\Phi_0$$

which is the EOM-CC eigenvalue problem for all excited states. Hence, once we have a MR-CC Ψ_g that correctly separates, *all* the excited states that come from the EOM-MRCC equations have to do the same, subject to certain symmetry conditions being satisfied by the MR-CC solution. For a closed shell reference in normal SR EOM-CC all excited states will naturally be spin-adapted.

Our recently published paper on EOM-CCx [15] is pertinent here. The x indicates an extension that bridges the gap between standard EOM-CC and VU-MR-CC that allows for the more correct treatment of other states like charge-transfer ones. The interplay between the two approaches pertains to the MR-DI/DA-CC methods, too. For a situation like $(AB)^{+2}\rightarrow A^{+2}+$ B or A + B^{+2} there is no problem, but if we also want to be able to explore channels like $(AB)^{+2}\rightarrow A^{+1}+\ B^{+1}$ we will have a similar, MR-DI-CCx method. Hence, such additions open the door to many other situations.

ACES III

As in all advances in the field, in addition to the development of the new methods



summarized above, the theory has to be incorporated into widely available and easily applied software to benefit its target audience. Furthermore, to treat the large molecules required, today it is mandatory to have the methods implemented into massively parallel programs that take full advantage of the extensive numbers of computational cores that

are now available, for example, at the DoD High Performance Computing Modernization Program (HPCMP) Supercomputing Resource Centers (SRC). ACES III was written from scratch under CHSSI support to make it possible to do very large scale CC applications for molecules. Consequenlty, an entirely different strategy was followed, including writing a new, domain specific language, SIAL, which means super-instruction assembler language, and a compiler for its commands [18]. In this way, mostly intuitive instructions written by quantum chemistry programmers generate operations called super-instructions that process very large numbers as a block instead of individually. The super instructions are executed by the super instruction processor (SIP), which takes care of all intricate details related to efficient, massively parallel excecution of the SIAL program. This effectively separates the quantum chemical program from the underlying message passing (MPI) and memory handling that ultimately makes the program perform exceptionally well. Furthermore, this separation enables the quantum chemistry programmer to rapidly add new capabilities and fine-tune others without having to touch the underlying SIP structure.

Some numerical examples are shown in Figs 8-12. The first is for MP2=MBPT2 analytical gradient calculations which, along with the MP2 energies, will often provide the first-cut at the PES for molecules. At this inexpensive level, we also have hessians (analytical second derivatives) to document the nature of the structure and transition states for a molecule and to provide its vibrational spectra. For RDX with 21 atoms, 114 electrons and 1005 basis functions, the scale runs from 100 to about a minute as a function of 100 to 4000 cores. For the larger HMX molecule (28 atoms, 152 electrons and 1340 basis functions the comparative performance is shown.

The next level is CCSD which is a far more demanding calculation. This figure shows that excellent scaling is achieved up to 4000 processors for RDX and about 6000 for the larger HMX. (We now use 924 functions in the latter case to achieve rapid turnaround for timing purposes). The bottom line is that such a calculation should take no more than an hour given adequate numbers of cores.

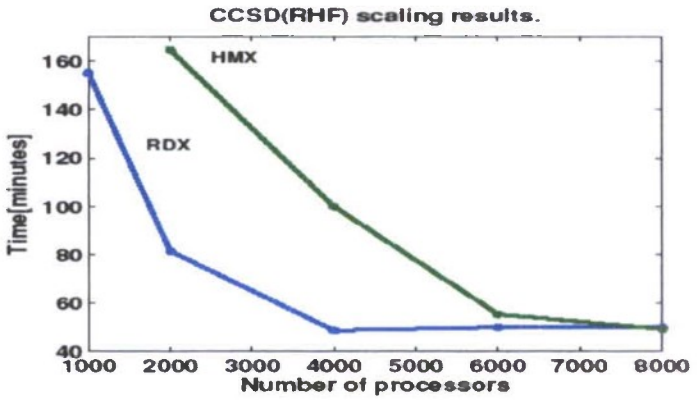


Fig. 8

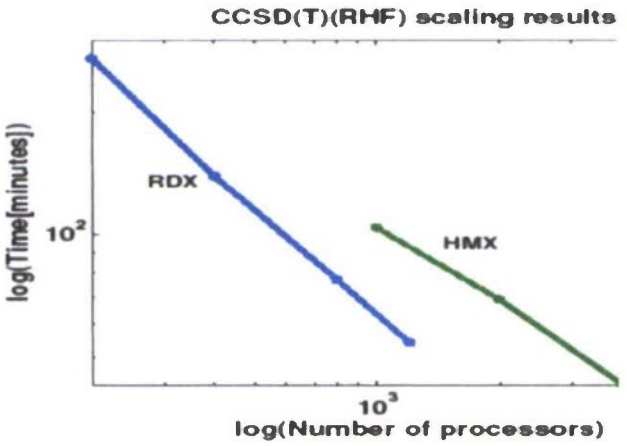
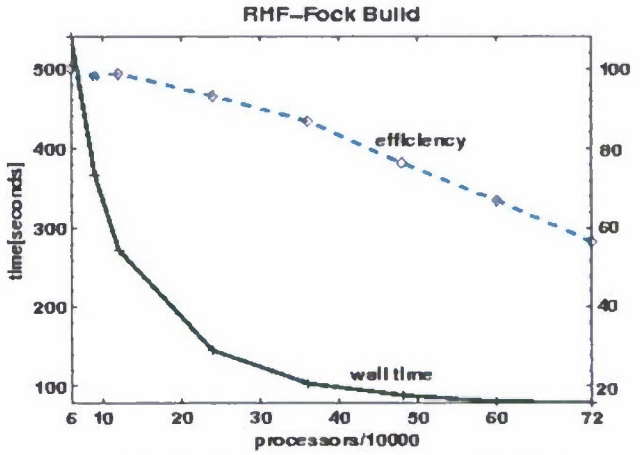


Fig. 9

Since CCSD scales ~ n^2N^4, when adding the (T) part to obtain CCSD(T) the scaling increases to ~ n^3N^4. This increase makes such calculations very time-consuming, but because of the comparative lack of IO and other kinds of processing, the (T) part scales very well in parallel. Here we see that once we hit 1000 cores RDX with 372 functions and HMX with 496 (T) takes ~ 100 minutes, but can be substantially further reduced as the number of cores is increased. The reduced basis was again chosen for comparztive timing purposes, but we are routinely applying (T) with > 1000 basis functions in our scientific applications.

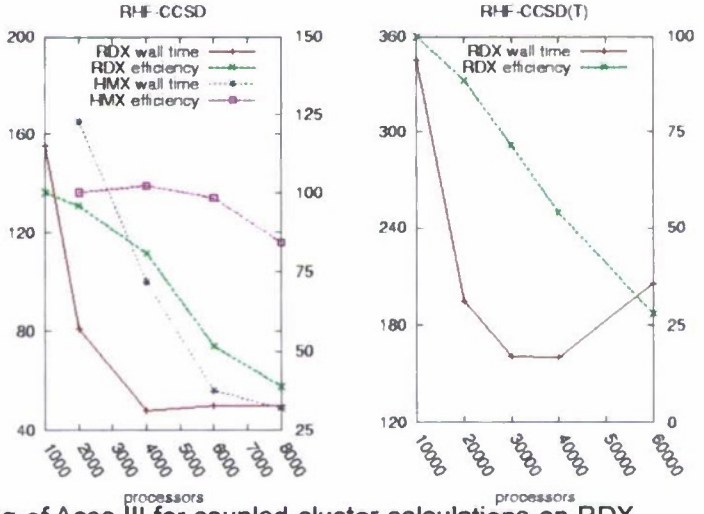
The next figure that shows the time for the Fock build for a nano–diamond with a N impurity demonstrates that we can readily employ as many as 75,000 cores for the Cray XT5 system, similar to the new Cray XE6now installed at the HPCMP SRC installations.

The final figure compares efficiency between calculations on RDX and HMX on many-thousands of cores.



Diamond nano crystal with Nitrogen vacancy C42H42N+ in an aug-cc-pvtz basis of 2944 functions. The graph shows the wall time in seconds and the efficiency for 6,000, 9,000, 12,000, 24,000, 36,000,48,000, 60,000, 72,000 for building the Fock matrix.

Fig. 10



Scaling of Aces III for coupled-cluster calculations on RDX (C3H6N6O6) and HMX (C4H8N8O8) on the Cray XT5 system at ORNL (jaguar). The left vertical axis of each figure is the wall time in minutes while the right axis is the scaling efficiency with respect to the first data point.

Fig. 11

Another example of the current performance of ACES III is the successful application of the EOM-CCSD method for the excited states of the cytosine molecule and its derivatives. To mimic the biological environment we have added one, two and four water molecules to cytosine, as well as a sugar ligand. By this we arrived to a system size with almost 100 correlated electrons which before ACES III were not possible to treat at this level of theory. The table shows that even for the largest system as much as 12 excited states could be calculated within a day with even modest numbers of cores. For the largest problem, each state takes an average of 2 hrs elapsed time.

This example shows clearly that the EOM method can be used very efficiently on large-scale computers, and therefore one can expect that our MR methods MR-DI-CCSD and MR-DA-CCSD will run with outstanding efficiency by using the technology we presently have in ACES III. Also, one of our Phase II objectives is to further develop EOM-MR-CC to ensure that the correct separation achieved by the MR method in the ground state will permit accurate asymptotes for all the subsequent excited states obtained. This development will also build upon the current EOM-CC methodology in ACES III.

EOM-CCSD/aug-cc-pVDZ calculations of 12 excited states.

	Cytosine	Cytosine + 1 water	Cytosine +2 waters	Cytosine + 4 waters	Cytidine
# atoms:	13	16	19	25	30
# correlated electrons:	42	50	58	74	94
# bfs:	229	270	311	393	508
Elapsed time	3.5 hours	4.5 hours	8.5 hours	21.9 hours	24 hours
Cores used	128	128	256	256	512

Fig. 12

The fact that ACES III provides the fastest current implementation of SR-CC that provides everything required: analytical gradients, density matrices, excited, ionized, and electron attached states, demonstrates that our framework for massively parallel processing is exceptional. Now from Phase I, ACES III already has TCCSD(T) and MR-

DI-CCSD in it. As we add analytical gradients to these and further develop the @CC approach building upon SIAL and the SIP processes, it guarantees that the new MR methods that will be implemented in Phase II will more than maintain this exceptional standard.

For comparison purposes, the serial ACES II contains the MR methods, SU-CCSD, SS-MkCC, and SS-BW-CC. This makes it convenient to provide definitive comparison results to test and further develop the new methods we are proposing. This capability will also facilitate adding the older, more conventional MR-CC methods to ACES Iii if proven worthy.

References

1. R. J. Bartlett and M. Musial, "Coupled-cluster theory in quantum chemistry", Revs. of Modern Phys. **79,** 291-352 (2007).
2. R. J. Bartlett, "The Coupled-cluster Revolution," Mol Phys. **108** (21-23), 2905-2920 (2010).
3. R. J. Bartlett, "Coupled-cluster theory and its equation-of-motion extensions," Wiley Inter-Science Reviews, Ed. Walter Thiel, in press.
4. Using frozen natural orbitals (FNO)[A. Taube and R.J. Bartlett, "Improving upon CCSD(T):ΛCCSD(T). Potential energy surfaces," J. Chem. Phys. 128, 044110/1-044110/13 (2008).], decreases the size of v by about a factor of 2, saving 2^4 in CCSD and CCSD(T) and would save 2^5 in CCSDT and 2^6 in CCSDTQ. So that is an easy improvement. The methods CCSD(T), ΛCCSD(T) save over CCSDT-3 by bypassing the T3 iterations in the latter, a saving of about an order of magnitude, and CCSDT(Q_f) does the same compared to CCSDTQ-1.

   The next level of simplification is offered by selecting just a few 'active' orbitals to introduce the effects of T3 and T4. These variants are termed CCSDt and CCSDtq [P. Piecuch, S. A. Kucharski and R. J. Bartlett, "Coupled-cluster methods with internal and semi-internal triply and quadruply excited clusters: CCSDt and CCSDtq approaches," J. Chem. Phys. 110, 6103- 6122 (1999).]. The scaling now is o^2v^4 for the underlying CCSD, but adding t only costs $o^2v^4o_Av_A$, and q, $o^2v^4(o_Av_A)^2$. Considering that the number of active orbitals is likely to be only a handful compared to hundreds for a normal, large scale application today; this is an enormous saving over the full CCSDTQ method. At this level, which is the full CI for 4 electrons and all *products* of four electron units, we would expect to be able to describe any MR effects that would involve four electrons and four orbitals, a level not yet reached by the operable MR-CC methods. If the choice of the active orbitals is good, and again FNO's would seem to be maybe the best choice, then CCSDtq should well approximate CCSDTQ and be applicable at a reasonable cost. See Phase II plan.

   There are also simplifications possible by exploiting localized orbitals. that can even achieve linear scaling with careful application. See eg, C. Hampel and H.-J. Werner, J. Chem. Phys. 104 6286 (1996); M. Schütz and H.-J. Werner, J. Chem. Phys. 114 661 (2001).
5. M. A. Haque and D. Mukherjee, "Application of cluster expansion techniques to open-shells: calculation of difference energies," J. Chem. Phys. 80, 058-5069 (1984).
6. I.Lindgren, "A coupled-cluster approach to the many-body perturbation theory for open-shell systems," Int. J. Quantum Chem., Quantum Chem. Symp. 12, 33-58 (1978).

7. A. Balkova and R. J. Bartlett, "A multireference coupled-cluster study of the ground state and lowest excited states of cyclobutadiene," J. Chem. Phys. 101, 8972-8987 (1994).
8. X. Li and J. Paldus, "The general-model-space state-universal coupled-cluster method exemplified by the LiH molecule," J. Chem. Phys. 119, 5346-5357 (2003).
9. I.Hubač, J. Pittner, and P. Čársky, "Size-extensivity correcetion for the state-specific multireference Brillouin-Wigner coupled-cluster theory," J. Chem. Phys. 112, 8779-8784 (2000).
10. U.S. Mahapatra, B. Datta, and D. Mukherjee, "A state-specific multi-reference coupled cluster formalism with molecular applications," Mol. Phys. 94, 157-171 (1998).
11. D.Lyakh, V. Lotrich, R. J. Bartlett, "Multireference world of chemistry," Chem. Rev., to be published.
12. D. Lyakh, V. Lotrich, R. J. Bartlett, "The "tailored" CCSD(T) description of the automerization of cyclobutadiene ",Chem. Phys. Letts. **501**, 166-171 (2011).
13. M. Musial, A. J. Perera, R. J. Bartlett, "Multi-reference Coupled-cluster Theory: The Easy Way," J. Chem. Phys., in press.
14. D. Lyakh, R. J. Bartlett, "Towards an Adaptive Coupled Cluster Theory," J. Chem. Phys. **133** 244112 (2010).
15. M. Musial, R. J. Bartlett, "Charge-Transfer Separability and Size-Extensivity in the Equation-Of-Motion Coupled Cluster Method: EOM-CCx ", J. Chem. Phys., **134**, 034106 (2011).
16. M. Musial and R.J. Bartlett, "Multi-reference Fock space coupled-cluster method in the intermediate Hamiltonian formulation for potential energy surfaces," J. Chem. Phys., to be published.
17. M.R. Berman, T. Tsuchiya, A. Gregusova, S.A. Perera and R.J. Bartlett, "HNNC radical and its role in the CH + N_2 reaction," J. Phys. Chem. A 111, 6894-6899 (2007).
18. V.F. Lotrich, J.M. Ponton, A.S. Perera, E. Deumens, R.J. Bartlett, and B.A. Sanders, "Super instruction architecture of petascale electronic structure software: the story," Mol. Phys. **108** (21-23), 3323-3330 (2010).